GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2003, 18:23:55 / Search time 2374.81 Seconds
                 (without alignments)
                 16364.623 Million cell updates/sec

Title: US-10-026-106E-7
Perfect Score: 1599
Sequence: 1 aaggccatggcgggggcccga.........acatcccaacgacctgatg 1599

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 1215223805b residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :

```
1: em_estba:*
2: em_esthum:*
3: em_estmu:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: em_esti:*
10: gb_est2:*
11: gb_hte:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estcm:*
16: em_estcm:*
17: em_estfun:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_vrt:*
21: em_gss_pln:*
22: em_gss_fun:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	592	37.0	1078	12	BQ061188	BQ061188 AGENCOURT
2	490.6	30.7	508	10	BE244935	BE244935 TCBAP1D26
3	388.4	24.3	430	10	BE246345	BE246345 TCBAP1D29
c 4	383.4	24.0	836	13	BU151377	BU151377 AGENCOURT

c	5	381.8	23.9	991	12	BQ056204	BQ056204 AGENCOURT
c	6	361.6	22.6	476	28	AQ610898	AQ610898 HS_5105_A
c	7	359.8	22.5	964	13	BQ894374	BQ894374 AGENCOURT
c	8	348	21.8	541	28	AQ610868	AQ610868 HS_5105_A
9	338.2	21.2	358	13	BU430530	BU430530 UI-HF-BN0
10	330.8	20.7	640	14	BY729118	BY729118 RPCI-11-4
11	292	18.3	558	28	AQ634533	AQ634533 RPCI-11-4
12	227.2	14.2	620	12	BI104593	BI104593 602B91434
13	215.2	13.5	496	13	CB713061	CB713061 AMGNNUC:N
14	204	12.8	460	14	CB738101	CB738101 AMGNNUC:M
c	15	201.2	12.6	506	12	AQ772013	AQ772013 HS_5413_B
c	16	189.2	11.8	609	12	BI657668	BI657668 603281241
c	17	134.4	8.4	555	14	W16444	W16444 za98a12.r1
18	89	5.6	392	13	BY093564	BY093564 BY093564
19	56.2	3.5	663	12	BJ043372	BJ043372 BJ043372
20	51.2	3.2	800	13	BU237675	BU237675 603411248
c	21	50.2	3.1	925	29	CNS0091P	CNS0091P Drosophil
c	22	47.4	3.0	391	14	CB755730	CB755730 AMGNNUC:8
23	45	2.8	290	10	BF171796	BF171796 PCL3410 M
c	24	44.4	2.8	641	12	BE293269	BE293269 601143880
25	44.4	2.8	642	12	BM784359	BM784359 K-EST0062
c	26	44.4	2.8	824	12	BG821330	BG821330 602724958
c	27	44.4	2.8	852	10	BG468516	BG468516 602510492
c	28	44.4	2.8	968	13	BQ420551	BQ420551 AGENCOURT
c	29	44.4	2.8	1000	13	BQ066710	BQ066710 AGENCOURT
30	44.4	2.8	1201	13	BX360822	BX360822 BX360822
31	44.4	2.8	1201	13	BX421108	BX421108 BX421108
32	44.2	2.8	337	10	BF087448	BF087448 QV2-HT054
33	44	2.8	1122	12	BQ059215	BQ059215 AGENCOURT
34	44	2.8	1201	13	BX363899	BX363899 BX363899
35	44	2.8	1201	13	BX381071	BX381071 BX381071
36	43.8	2.7	341	29	B2367545	B2367545 1d05b01.b
37	43.4	2.7	1135	13	BQ278875	BQ278875 AGENCOURT
38	43.4	2.7	280	12	BI063605	BI063605 PM4-UT010
c	39	43.2	2.7	388	12	BI199792	BI199792 602761012
40	43	2.7	709	28	B2121065	B2121065 CH230-317
41	43	2.7	837	29	CNS02GIU	CNS02GIU Tetraodon
42	43	2.7	1101	29	CNS017SY	CNS017SY Drosophil
43	42.8	2.7	878	13	BU911485	BU911485 Drosophil
44	42.8	2.7	514	13	BY008068	BY008068 AGENCOURT
45	42.4	2.7	967	12	BI108915	BI108915 602896666

ALIGNMENTS

RESULT 1
BQ061188
LOCUS   BQ061188   1078 bp   mRNA   linear   EST 02-APR-2002
DEFINITION AGENCOURT_6863006 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5920208
            5', mRNA sequence.
ACCESSION   BQ061188
VERSION   BQ061188.1 GI:19884982
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1078)
AUTHORS   NIH-MGC http://mgc.ncl.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt.
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2083 row: 1 column: 09
            High quality sequence stop: 593.

FEATURES   Location/Qualifiers
source   1. .1078
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /clone="IMAGE:5920208"
                  /db_xref="taxon:9606"
                  /tissue_type="lymphoma, cell line"
                  /lab_host="NIH_MGC_99"
                  /clone_lib="DH10B (phage-resistant)"
                  /note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
                  EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                  into EcoRI/XhoI sites using the following 5' adaptor:
                  GGCACGAG(G). Size-selected >500bp for average insert size
                  1.8kb. Library constructed by Ling Hong in the laboratory
                  of Gerald M. Rubin (University of California, Berkeley)
                  using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                  II RT (Life Technologies). Note: this is a NIH_MGC
                  Library."

BASE COUNT   243 a   306 c   309 g   219 t   1 others
ORIGIN

Query Match   37.0%;   Score 592; DB 12; Length 1078;
Best Local Similarity   97.9%; Pred. No. 2.4e-126;
Matches 619; Conservative 0; Mismatches 10; Indels 3; Gaps 2;

Qy   63 AGGGAGGGCCCCGGTCTCGGCCCCTCCCCAGAATGTGACGGTCGCTTCTCCCAGAACTTCAGCGT 122
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   157 AAGGAGGGCCCCGGTCTCGGCCCCTCCCCAGAATGTGACGGTCGCTCTCCCAGAACTTCAGCGT 216

Qy   123 GTACCTGACAGCATGGCTTCCCAGGGGCTTGGCCAACCCCCAGCCAGGATGTGACCTATTTTTGTGGCC-A 181
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   217 GTACCTGACAGCATGGCTTCCCAGGGGCTTGGCCAACCCCCAGGATGTGACCTATTTTTGTGGCCTA 276

Qy   182 TCAGAGGCTCTCCCACCCGTAGACGGTGGCGCCGAAGTGGAAGAGTGTGCGCGGAACCAAGGA 241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   277 TCAGAGGCTCTCCCACCCGTAGACGGTGGCGCCGAAGTGGAAGAGTGTGCGGGAACCAAGGA 336

Qy   242 GCTGCTATGTTCTATGATGTGCGGAAGAAACAAGGACCTGTACAACAAGTTCAAGGGACG 301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   337 GCTGCTATGTTCTATGATGTGCCTGAAAGAAACAGGACCTGTACAACAAGTTCAAGGGACG 396

Qy   302 CGTGCGGAGGAGTTTCTCCCAGCTCCAAGTCCCCTCTGGGTGGAGTCGAATACCTGGATTA 361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   397 CGTGCGGAGGAGTTTCTCCCAGCTCCAAGTCCCCTCTGGGTGGAGTCCGAATACCTGGATTA 456

Qy   362 CCTTTTTGAAGTGGAGCGGCCGGCCCCACCTGTCTCCTGGTGCTCACCCAGACGGAGGAGATCCT 421
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   457 CCTTTTTGAAGTGGAGCGGCCGGCCCACCTGTCTCCTGGTGCTCACCCAGACGGAGGAGATCCT 516

Qy   422 GAGTGCCGAATGGCCACGTACCAGTCGTGCCCCTTGCATGGCATCTGGATCTGAAGTATGA 481
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   517 GAGTGCCGAATGGCCACGTACCAGTCGTGCCCCTCGCATGGCATCTGGATCTGAAGTATGA 576

Qy   482 GGTGGCATTCTGGAAGGAGGGGGCCCGGAAACAAGACCCTATTTCCAGTGCACTCCCCATGG 541
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   577 GGTGGCATTCTGGAAGGAGGGGGCCGGAAACAAGACCCTATTTCCAGTGCACTCCCCATGG 636

Qy   542 CCAGCCAGTCCAGATCACTTCTCAGACATCCAGCCAGCTGCCTGCCTCAGTGCCAG 601
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   637 CCAGCCAGTCAGATCACTCTCAGACACCACTCCTGCCTCAGTGCCAG 696

Qy   602 AACCACTTCACGTTCAGTGTGTCCCGAAATACAGCAAGTTCTCTAGGCCACCACTGCTTCTT 661
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   697 AACCATCTACACGTTCAGTGTCCCGAAATACAGCAAGTTCTCTAAGCCACCACTGCTTCNT 756

Qy   662 --GCTGGAGGTCCCAGAAGGGAACTGGGCTTT 691
            ||||||||||||||||||||||||||||||
Db   757 TGCTGGAGGTCCCCAGAAGGGAACTGGGGCTTT 788

RESULT 2
BE244935
LOCUS   BE244935   508 bp   mRNA   linear   EST 03-OCT-2001

DEFINITION   TCBAP1D2669 Pediatric pre-B cell acute lymphoblastic leukemia
             Baylor-HGSC project=TCBA Homo sapiens cDNA clone TCBAP2669, mRNA
             sequence.
ACCESSION   BE244935
VERSION   BE244935.1 GI:9096765
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 508)
AUTHORS   Wei,Y.; Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman Jr.,P.R., Muzny,D.
             Bouck,J.; Gibbs,R.A. and Margolin,J.F.
TITLE   Pediatric Leukemia cDNA Sequencing Project
JOURNAL   Unpublished
COMMENT   Contact: Dr. Judith F. Margolin
             Texas Children's Cancer Center and Human Genome Sequencing Center
             at Baylor College of Medicine
             1102 Bates, MC3-3320 Houston, TX 77030, USA
             Tel: 832-824-4536
             Fax: 832-825-4038
             Email: clones@txccc.org
             Citation: Carinci,P. and Hayashizaki,Y. High efficiency
             full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
             Seq primer: M13 primer.
FEATURES   Location/Qualifiers
source   1. .508
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="TCBAP2669"
             /sex="male"
             /tissue_type="leukopheresis"
             /cell_type="pre-B cell"
             /dev_stage="pediatric 2 years"
             /clone_lib="Pediatric pre-B cell acute lymphoblastic
             leukemia Baylor-HGSC project=TCBA"
             /lab_host="DH10B"
             /note="Vector: lambda psB; Site_1: BamHI; Site_2: EcoRI;
             First strand cDNA was primed with an anchored
             XhoI-oligo(dT) primer [5'GGAGACTCGAGCGGCCGCCAGGAGGAG(T)VN
             3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand
             was primed with a BamHI-dC primer
             [5'AGGAGCTCCGATCCGCGCCCCAATATAATPAAT(C) 3'].
             Double-stranded cDNA was then digested with BamHI and XhoI
             and directionally cloned into the BamHI and SalI sites of
             lambda psB vector. Library went through one round of
             normalization. Library was constructed by Wei Yu at RIKEN
             of Japan (Carinci P, Westover A, Nishiyama Y, Ohsumi T,
             Itoh M, Nagaoka S, Sasakin, Okazaki Y, Muramatsu M,
             Schneider C, Hayashizaki Y, High efficiency selection of
             full-length cDNA by improved biotinylated cap trapper.,
             DNA Res 4: 1, 61-6, Feb 28, 1997)"

BASE COUNT   108 a   147 c   160 g   93 t
ORIGIN

Query Match   30.7%;   Score 490.6; DB 10; Length 508;
Best Local Similarity   99.0%; Pred. No. 5.4e-103;
Matches 504; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy   928 TCAGGGCCCCAGCGCCACCCCCACACAGACAGAGATGGAAGAAGGACCTTGCAGAGAGGACGAAGAGG 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TCAGGGCCCCAGCGCCCCCACACAGACAGAGATGGAAGAAGGACCTTGCAGAGAGGACGAAGAGG 60

Qy   988 AGGAGGATGAGGAGGACCACAGAGAGATGGCGTCAGCTTCCAGCGGCCTACATTGAACCACTTT 1047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AGGAGGATGAGGAGGACCACAGAGAGATGGGCGTCAGCTTCCAGCGGCCTC-CATTGAACCGTTTT 119

Qy 1048 CTTTCCTGGGGCAAGAGCACCAGGTCTCCAGGGCACTCGGAGCCCGAGGCCTGGTTGTCGGGGACTCAG 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   120 CTTTCCTGGGGCAAGAGCACCAGGTTCCAGGGTACTCGGAGCCCTGGTTGTGGGGGTGGGGGACTCAG 179

Qy 1108 GGAGGCCCCAGGGGCTCCTCGGTCCCAAGCCAAGGCTCCTCGTCTTGGGGATTCTTCAGACA 1167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   180 GGAGGCCCCAGGGGCTCCTCGGTCCCAAGCCAAGGCTCCTCGTGGTGTTGGGGATTCTTCAGACA 1167

RESULT 3

```
Qy    180 GGAGGCCCAGGGTCCTCTGGTCCCAAAGCGAAAGGCGTCCTTCGTTCGCTTGGGATTCTTCAGACA 239
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1168 GAAGCTGGGCGCCAGCACTGTGGGACTCCTTCCTGGGACAGGGCGTCCTCGGCTATTTGG 1227

Qy   1168 GAAGCTGGGCGCCAGCACTGTGGGACTCCTCCTGGGACAGGGCGTCCTCGGCTATTTGG 1227
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 GAAGCTGGGCGCCAGCACTGTGGGACTCCTCCTGGGACAGGGCGTCCTCGGCTATTTGG 299

Qy   1228 CTGAGAAGGGGCCCAGGGCCAGGGGCCACCAAGAGAATCTCTCCCACCAC 1287
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 CTGAGAAGGGGCCCAGGGCCGGGGCCACCAAGAGAATCTCTCCCACCAC 359

Qy   1288 CTGAATTCTCCAAGGACTCGGGTTTCCTCTGGAAGAGCTCCCAGAAGATAACCTCTCCTCCTCCT 1347
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 CTGAATTCTCCAAGGACTCGGGTTTCCTGGAAGAGCTCCCAGAAGATAACCTCTCCTCCT 419

Qy   1348 GGGGCCACCTGGGGGCCACCTTACCACCGGGAGCGCGAATCTGGTCTGTCCCTGGGGGGACCCCAGTTT 1407
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 GGGGCCACCTGGGGGCCACCTTACCACCGGGAGCGCGAATCTGGTCTGTCCCTGGGGGGACCCCAGTTT 479

Qy   1408 CTCTTCAGACACTGACCTTCTGCTGGGAA 1436
          |||||||||||||||||||||||||||||
Db    480 CTCTTCAGACACTGACCTTCTGCTGGGAA 508
```

LOCUS   BE246345   430 bp   mRNA   linear   EST 03-OCT-2001
DEFINITION TCBAP1D2918 Pediatric pre-B cell acute lymphoblastic leukemia
            Baylor-HGSC project=TCBA Homo sapiens cDNA clone TCBAP2918, mRNA
            sequence.
ACCESSION   BE246345
VERSION   BE246345.1 GI:9098094
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 430)
AUTHORS   Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman Jr.,F.R., Muzny,D.
            , Bouck,J., Gibbs,R.A. and Margolin,J.F.
TITLE   Pediatric leukemia cDNA Sequencing Project
JOURNAL   Unpublished
COMMENT   Contact: Dr. Judith F. Margolin
            Texas Children's Cancer Center and Human Genome Sequencing Center
            at Baylor College of Medicine
            1102 Bates, MC3-3320 Houston, TX 77030, USA
            Tel: 832-824-4536
            Fax: 832-825-4038
            Email: clones@txccc.org
            Citation: Carnini,P. and Hayashizaki,Y. High efficiency
            full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
            Seq primer: M13 primer.
FEATURES   Location/Qualifiers
     source   1..430
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TCBAP2918"
                     /sex="male"
                     /tissue_type="leukopheresis"
                     /cell_type="pre-B cell"
                     /dev_stage="pediatric 2 years"
                     /clone_lib="Pediatric pre-B cell acute lymphoblastic
                     leukemia Baylor-HGSC project=TCBA"
                     /note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
                     First strand cDNA was primed with an anchored
                     Xhol-oligo(dT) primer (5'GGAGGACTCGAGCGGCCGCAGGAGGAG(T)VN
                     3'; V=A,C,G; N=A,C,G,T) and then dG tailed. Second strand
                     was primed with a BamHI-dC primer
                     (5'AGAGAGCTCGGATCCGCGCGGCCGCAATAAATAAT(C) 3').
                     /lab_host="DH10B"
                     Double-stranded cDNA was then digested with BamHI and XhoI
                     and directionally cloned into the BamHI and SalI sites of
```

lambda pSB vector. Library went through one round of
normalization. Library was constructed by Wei Yu at RIKEN
of Japan (Carnini P., Westover A., Nishiyama Y., Ohumi T,
Itoh M, Nagaoka S, Sasaki N, Okazaki Y, Muramatsu M,
Schneider C, Hayashizaki Y, High efficiency selection of
full-length cDNA by improved biotinylated cap trapper.,
DNA Res 4: 1, 61-6, Feb 28, 1997)"

```
BASE COUNT    95 a   120 c   136 g   79 t
ORIGIN
```

Query Match   24.3%; Score 388.4; DB 10; Length 430;
Best Local Similarity   97.0%; Pred. No. 2.3e-79;
Matches 417; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

```
Qy    928 TCAGGGCCCCAGCCACCCAAACACACACAAGAGATGGAAGAAGGACCCTTGCAGAGGACGAAGAGG 987
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TCAGGGCCCCAGCCACCCAAACACACACAAGAGATGGAAGAAGGACCCTTGCAGAGGACGAAGAGG 60

Qy    988 AGGAGGATGAGGAGGACACACAGAAGATGGCGGTCAGCTTCAGCCCTACATTGAACCACCTT 1047
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGGAGGATGAGGAGGACACACAGAAGATAGCGTCAGCTTCCAGCTCTTCTTTGAACCACGGTT 120

Qy   1048 CTTTTCCTGGGGCAAGAGCACCAGGTTCCAGGCACTCGGAGGCTTGGTGGGTGGGACTCAG 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTTTTCCTGGGGCCAAGAGCACCAGGCTCTATTGCACTCGGAGGGTTGGTGGGTGGACTCAG 180

Qy   1108 GGGAGGCCCAGGGCTTCCTCCTTCCTGGTCCCAAGGCAAGAGGAGGCTCCTCCTGGTTGGGATTCTTCAGACA 1167
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGGAGGCCCAGGGCTTCCTCCTTCCTGGTCCCAAGCAAGAGGAGGCTCCTCCTGCTTGGGGATTCTTCAGACA 240

Qy   1168 GAAGCTGGGCGCCAGCACTGTGGACTCCTTCCTGGGACAGGGCGTCCTCGGCTATTTGG 1227
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GAAGCTGGGCGCCAGCACTGTGGGACTCCTCCTGGGACAGGGCGTCCTCGGCTATTTGG 300

Qy   1228 CTGAGAAGGGGCCCAGGGCCAGGGGCCGGGGTGGGGATGGGACCACCAAGAGAATCTCTCCCACCAC 1287
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTGAGAAGGGGCCCAGGGCCAGGGCTGGGGGTGGGGATGGGGCACCAAGAGAATCTCTCCCACCAC 360

Qy   1288 CTGAATTCTCCAAGGACTCGGG-TTTCCTGGAAGAGCTCCCAGAAGATAACC-TCTCCTC 1345
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||| |||||||
Db    361 CTGAATTCTCCAAGGACTCGGGTTTCCTGGAAGAGCTCCCAGAAGATAAACCTTCTCCTC 420

Qy   1346 CTGGGGCCACC 1355
          |||||||||||
Db    421 CTGGGGCCACC 430
```

RESULT 4

LOCUS   BU151377
            BU151377/c
DEFINITION AGENCOURT_8585742 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:6304649
            5', mRNA sequence.
ACCESSION   BU151377   836 bp   mRNA   linear   EST 03-SEP-2002
VERSION   BU151377.1 GI:22664909
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 836)
AUTHORS   NIH-MGC http://mgc.ncl.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2525 row: k column: 18

RESULT 5

LOCUS   BQ056204/c
DEFINITION AGENCOURT_6773374 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5808505
            5', mRNA sequence.
ACCESSION   BQ056204
VERSION   BQ056204.1 GI:19815544
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 991)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov

LOCUS   BQ056204   991 bp   mRNA   linear   EST 29-MAR-2002
DEFINITION AGENCOURT_6773374 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5808505

FEATURES   Location/Qualifiers
     source   1..836
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6304649"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_99"
                     /note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."

High quality sequence stop: 522.

BASE COUNT   185 a   223 c   240 g   186 t   2 others
ORIGIN

Query Match   24.0%; Score 383.4; DB 13; Length 836;
Best Local Similarity   99.7%; Pred. No. 4.3e-78;
Matches   384; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   804 CTGGACTTTTCTGGACACACACACCCTGTGGCAACCTTTCAGCCCAGCAGACCAGAGTCC 863
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CAGGACTTTTCTGGACACACACCCTGTGGCAACCTTTCAGCCCAGCAGACCAGAGTCC 482

Qy   864 GTGAATGACTTGTTCCTCTGTCCCAAAAGGAACTGACCAGAGGGGGTCAGGCCGACGCCT 923
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GTGAATGACTTGTTCCTCTGTCCCAAAAGGAACTGACCAGAGGGGGTCAGGCCGACGCCT 422

Qy   924 CGAGTCAGGGCCCCAGCCACCCAACAGCAAGATGGAAGAAGAAGGACCTTGCAGAGGACGAA 983
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CGAGTCAGGGCCCCAGCCACCCAACAGCAAGATGGAAGAAGAAGGACCTTGCAGAGGACGAA 362

Qy   984 GAGGAGGAGGATGAGGAGGACACAGAGATGGCGTCAGCTTCCAGCCCTACATTGAACCA 1043
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GAGGAGGAGGATGAGGAGGACACAGAGATGGCGTCAGCTTCCAGCCCTACATTGAACCA 302

Qy   1044 CCTTCTTTCCTGGGGCCAAGAGACGCTCCAGGGCACTCGGAGGCTGGTGTGGGGTGGGAC 1103
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCTTCTTTCCTGGGGCCAAGAGACGCTCCAGGGCACTCGGAGGCTGGTGTGGGGTGGGAC 242

Qy   1104 TCAGGGAGGAGCGGCCCCAGGGGCTTCCTCTGGGTCCCAAGGAAGGCTCCTCCGGATTCTTCTTCA 1163
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TCAGGGAGGAGCGGCCCCAGGGGCTTCCTCTGGGTCCCAAGGAAGGCTCCTCCGGATTCTTCTTCA 182

Qy   1164 GACAGAAGCTGGGCCCAGCACTGTGG 1188
            ||||||||||||||||||||||||||
Db   181 GACAGAAGCTGGGCCAGCACTGTGG 157

RESULT 6

LOCUS   AQ610898/c
DEFINITION HS_5105_A2_G06_SP6E RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=681 Col=12 Row=M; genomic survey sequence.
ACCESSION   AQ610898
VERSION   AQ610898.1 GI:5072174
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens

LOCUS   AQ610898   476 bp   DNA   linear   GSS 15-JUN-1999
DEFINITION HS_5105_A2_G06_SP6E RPCI-11 Human Male BAC Library Homo sapiens

FEATURES   Location/Qualifiers
     source   1..991
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5808505"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_99"
                     /note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."

High quality sequence stop: 645.

Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2052 row: k column: 02

BASE COUNT   224 a   285 c   270 g   211 t   1 others
ORIGIN

Query Match   23.9%; Score 381.8; DB 12; Length 991;
Best Local Similarity   99.5%; Pred. No. 1.1e-77;
Matches   383; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   804 CTGGACTTTTCTGGACACACACACCCTGTGGCAACCTTTCAGCCCAGCAGACCAGAGTCC 863
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CAGGACTTTTCTGGACACACACCCTGTGGCAACCTTTCAGCCCAGCAGACCAGAGTCC 482

Qy   864 GTGAATGACTTGTTCCTCTGTCCCAAAAGGAACTGACCAGAGGGGGTCAGGCCGACGCCT 923
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GTGAATGACTTGTTCCTCTGTCCCAAAAGGAACTGACCAGAGGGGGTCAGGCCGACGCCT 422

Qy   924 CGAGTCAGGGCCCCAGCCACCCAACAGCAAGATGGAAGAAGAAGGACCTTGCAGAGGACGAA 983
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CGAGTCAGGGCCCCAGCCACCCAACAGCAAGATGGAAGAAGAAGGACCTTGCAGAGGACGAA 362

Qy   984 GAGGAGGAGGATGAGGAGGACACAGAGATGGCGTCAGCTTCCAGCCCTACATTGAACCA 1043
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GAGGAGGAGGATGAGGAGGACACAGAGATGGCGTCAGCTTCCAGCCCTACATTGAACCA 302

Qy   924 CGAGTCAGGGCCCCAGCCACCCAACAGACAAGATGGAAGAAGAAGGACCTTGCAGAGGACGAA 983
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CGAGTCAGGGCCCCAGCCACCCAACAGCAAGATGGAAGAAGAAGGACCTTGCAGAGGACGAA 362

Qy   1044 CCTTCTTTCCTGGGGCCAAGAGACGCTCCAGGGCACTCGGAGGCTGGTGTGGGGTGGGAC 1103
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCTTCTTTCCTGGGGCCAAGAGACGCTCCAGGGCACTCGGAGGCTGGTGTGGGGTGGGAC 242

Qy   1104 TCAGGGAGGAGCGGCCCCAGGGGCTCCTCTGGGTCCCAAGGAAGGCTCCTCCGGATTCTTCTTCA 1163
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TCAGGGAGGAGCGGCCCCAGGGGCTCCTCTGGGTCCCAAGGAAGGCTCCTCCGGATTCTTCTTCA 182

Qy   1164 GACAGAAGCTGGGCCCAGCACTGTGG 1188
            ||||||||||||||||||||||||||
Db   181 GACAGAAGCTGGGCCAGCACTGTGG 157

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```
REFERENCE    1  (bases 1 to 476)
AUTHORS      Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
             Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
             Hood,L.
TITLE        Sequence-tagged connectors: A sequence approach to mapping and
             scanning the human genome
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE      99380589
PUBMED       10449764
COMMENT      Contact: Mahairas GG, Wallace JC, Hood L
             High Throughput Sequencing Center
             University of Washington
             401 Queen Anne Avenue North, Seattle, WA 98109, USA
             Tel: (206) 616-3618
             Fax: (206) 616-3887
             Email: jwallace@u.washington.edu
             Clones are derived from the human BAC library RPCI-11.
             Clones are derived from the human BAC library RPCI-11. For BAC
             library availability, please contact Pieter de Jong
             (pieter@deejong.med.buffalo.edu). Clones may be purchased from
             BACPAC Resources (http://bacpac.med.buffalo.edu/ordering.bac.htm)
             or from Resear h Genetics (info@oregen.com). BAC end Web Server:
             http://www.htcc.washington.edu
             Plate: 681  row: M  column: 12
             Seq primer: SP6
             Class: BAC ends
             High quality sequence stop: 476.
FEATURES             Location/Qualifiers
     source          1..476
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=681 Col=12 Row=M"
                     /sex="male"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
BASE COUNT      96 a    137 c    128 g    115 t
ORIGIN
```

```
Query Match                22.6%;  Score 361.6;  DB 28;  Length 476;
Best Local Similarity   89.9%;  Pred. No. 3.8e-73;
Matches   410;  Conservative   0;  Mismatches  44;  Indels  2;  Gaps  2;

Qy    834 GCAACCTTTCAGCCCAGCGAACCAGAGTCCGTGAATAACTGTTCTCTGTCCCCAAAAG 893
          ||||||||||||||||| | |||||||| | ||||||| || || |||| ||||||||
Db    454 GCTACCTTTCAGCCCAGCAGACCAGAGTCAGTGAATAGCTTGTTCCTCTCGTCCCCAAAAG 395

Qy    894 GAACTGACCAGAGGGGTTCAGGCAAGCGCTTGTGAGTCAGGGCCCCAGCCACCAACAGACA 953
          ||||||| ||||||||||||||||||||| || |||||||||||| |||| ||||||||||
Db    394 GAGCTGACCAGAGGGTTCAGGCAAGCGCTTGTGAGTCAGGGCCCCAGCCACCAACAGACA 335

Qy    954 AGATGGAAGAAGGAAGGACCTTGCAGAGGACGAAGAGGGATGGAGGAGGACCACAGAAGAT 1013
          ||||||||||||| |||||||||||||||||||||||||||||||||| ||||| ||||||
Db    334 AGATGAAAGAAGGTCCTT-CATAGTTCGAAGAGCAGTAGGAGTAGGAGTCACATATGAT 276

Qy   1014 GGCGGTCAGCTTCCAGCCCTACATTGAACCACCTTCTTTCCTTGTTCTCTGTCCCCAAGCT 1073
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    275 GGCGGTCAGCTTCCAGCCATACACCTTCTTTCGTGGGGGCAAGCAGCACCAGCT 216

Qy   1074 CCAGGGGCACTCGGAGGCTGCTGTGGGGGTGGACTCAGGGAGGGCCCAGGGCTCCTCTGGTCCCA 1133
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    215 CCAGGGCACTCGGAGGCTGGT-GGGTGAACTCAGGAGACCCAGTGCTCTTCTGGTCCCA 157

Qy   1134 AGGGAAGGCTCCTCTTGCTTGGGATTCTTCAGACAGGAGAAGCTGGGGCCAGCACTGTGGGACTCC 1193
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    156 AGAGAACGGCTCCTCTGCTTCGAATTCTTCAGACAGAAAACTGGGCCCCAGCACTGTGGGACTCC 97

Qy   1194 TCCTTGGGACAGGGCTGGGTCCCTCAGACTCATGTGGCTTAAGGAAGAAGGGGCCCAGGCCCAAGGGCCG 1253
```

```
RESULT 7
BQ894374/c
LOCUS       BQ894374              964 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_8623992 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:6304287
            5', mRNA sequence.
ACCESSION   BQ894374
VERSION     BQ894374.1  GI:22286388
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 964)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2524  row: 1  column: 16
            High quality sequence stop: 510.
FEATURES             Location/Qualifiers
     source          1..964
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6304287"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_99"
                     /note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
BASE COUNT      196 a    277 c    268 g    223 t
ORIGIN
```

```
Db     96 GGTGGGGGATGGGCACCAGGAGAATCTCCCCACCACCT 1289
          ||||||||||||||||||||||||||||||||||||||
Qy   1254 GGTGGGGGATGGGCACCAGGAGAATCTCCCCACCACCT 1289
Db     36 GGTGTGCATGTGCGCCAAGAATCTCCCCACCACAT 1
```

```
Query Match                22.5%;  Score 359.8;  DB 13;  Length 964;
Best Local Similarity   99.0%;  Pred. No. 1.3e-72;
Matches   383;  Conservative   0;  Mismatches   2;  Indels  2;  Gaps  2;

Qy    804 CTGGACTTTCTGGACACACACCACACCCTGTGGCAACCTTTCAGCCCAGCGAACCAGAGT-C 862
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    557 CAGGACTTTCTGGACACGCACCACACCCTGTGGCAACCTTTCAGCCCAGCGAACCAGAGTCC 498

Qy    863 CGTGAATGACTGTTCTCTGTCCCCAAAAGGAACTGACCAGAGGGTCAGGGGTCAGGCC-GACGGC 921
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    497 CGTGAATGACTGTTCTCTGTCCCCAAAAGGAACTGACCAGAGGGTCAGGGCCGGACGGC 438

Qy    922 CTCGAGTCAGGGGCCCCAGCCACCAACAGACAAGATGGAAGAAGGAAGGACCTTGCAGAGGACG 981
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    437 CTCGAGTCAGGGGCCCCAGCCACCAACAGACAAGATGGAAGAAGGAAGGACCTTGCAGAGGACG 378

Qy    982 AAGAGGAGGAGGATGGAGGAGGACAGAGGAGGACGAAGAGGCGGTCAGCTTCCAGCCCTACATTGAAC 1041
```

RESULT 8

```
Qy  377 AAGAGGAGGAGGATGAGGAGGACAGAGAAGATGGCGGTGGTCAGCTTCAGCCTACATTGAAC 318
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1042 CACCTTCTTCCTGGGGGGGCAAAGAGCACCAGGCCACTCGGAGGCTGGTGGTGGGGTGG 1101
Qy  317 CACCTTCTTCCTCGTGGGGCAAGAGCACCAGGCCTCCGAGGCTCGGAGGCTGGTGGTGG 258
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1102 ACTCAGGGAGGCCCAGGGGTCTCCTCGTTCCCAAGGGGAAGGCCTCTCTGGTTGGGATTCTT 1161
Qy  257 ACTCAGGGAGGCGCCCAGGGGTCTCCTCGTTCCCAAGGCGGAAGGCCTCTCTGGTTGGGATTCTT 198
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1162 CAGACAGGAAGCTGGGCGCCAGCACTGTGG 1188
Qy  197 CAGACAGAGAAGCTGGGCGCCAGCACTGTGG 171
```

```
LOCUS       AQ610868/c             541 bp    DNA    linear    GSS 15-JUN-1999
DEFINITION  HS_5105_A2_D06_SP6E RPCI-11 Human Male BAC library Homo sapiens
            genomic clone Plate=681 Col=12 Row=G, genomic survey sequence.
ACCESSION   AQ610868
VERSION     AQ610868.1  GI:5072144
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 541)
AUTHORS     Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
TITLE       Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE     99380589
PUBMED      10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@bacpac.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
FEATURES             Location/Qualifiers
     source          1..541
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="plate=681 Col=12 Row=G"
                     /sex="male"
                     /clone_lib="RPCI-11 Human Male BAC library"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
                     High quality sequence stop: 541.
                     Seq primer: SP6
                     Plate: 681  row: G  column: 12
                     Seq primer: SP6
                     Class: BAC ends
BASE COUNT    173 a    140 g    125 t    100 c    3 others
ORIGIN
Query Match     21.8%;   Score 348;  DB 28;  Length 541;
Best Local Similarity   87.3%;   Pred. No. 5.6e-70;
Matches  404; Conservative  0; Mismatches  56; Indels  3; Gaps  2;
```

RESULT 9

```
Qy  821 CACACACCCTGTGGCAAACCTTCAGCCCCAGCCAGAGACCAGAGTCCGTGGAATGACTTGTTCCT 880
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  532 CACACGCACCCTGTGCAACGTTTCAGCCAGCAGAGCCAGCCAGAGT-CGGGAATGACTTGTTCCT 474
Qy  881 CTGTCCCCAAAAGGAACTGACCACCAGAGGGGTCAGGCCGACGGCCTCGAGTCAGGGCCCCAGC 940
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  473 CTGTCCCAAAA--GAAACTGCCAGAGGGGTCAGGCCCGACGGCCTCGAGTCAGGGCCCCAGC 416
Qy  941 CACCCAACAGACACAAGATGTGGAAGAAGGAAGAACCTTTGCAGAGGAGGAGGAGGAGGATGAGGA 1000
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  415 CACACAACAGAGCAGGGAAGAGAGATGGAAGAAGAGGTTCCTTGCAGAGGAGGAGGAGGATGAGGA 356
Qy 1001 GGACACAGAGAAGATGGCGTCAGCTTCCAGCCCTTCCAGCCCTACATTGAACCGGCCCTTCTTCTTCCTGGGGGCA 1060
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  355 GTACACANAAGATGGCGTCAGCTTCCAGCCCTTCCAGCCCTACATTGAACCGGCCCTTCTTCTTCCTGGGGGCA 296
Qy 1061 AGAGCACCAGGCCTCAGGGCAGCTCCGGGAGGCTGGTGGTGGGGTGGGCCCAGGGC 1120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  295 AGAGCACCAGGCTACAGGGCCACTCGGAGACTCGGTGGTGGGGTGGGCCACAGGGC 236
Qy 1121 TCCTCTGGTTCCCAACCGAAGGGCTTCCTCTTGCTTGGGATTCTTCAGACAGCAGAAGCTGGGCGCCCAG 1180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  235 TCCGTCTGGTTCCCAAGGAGAGAGGCTCCTCTTGCTTGGGATTCTTCAGACAGCAGAAGCTGGGCGCCAG 176
Qy 1181 CACTGTGGACTTCCCTCCTGGGGACAGGGGCTGGGCTTTGGGCTATTTGGGCTGAGAAGCAGGGGGCC 1240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  175 CGCTGTGGACTTCCCATCCGGAACAGGAGGTGGGCCTCTATTTGGGCTGAGAAGCAGGGGAGC 116
Qy 1241 AGGCCAAGGGCCGCGGGTGGGGATGGCGGCCACCAAGAAATCTCTCCCA 1283
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db  115 AGGCCAAGGGCCGCGGGTGGTTATGGGCCCCAAGATCTTCCCCCACA 73
```

```
LOCUS       BU430530               358 bp    mRNA    linear    EST 09-SEP-2002
DEFINITION  UI-HF-BN0-aez-h-06-0-UI.r1 NIH_MGC_50 Homo sapiens cDNA clone
            IMAGE:3066106 5', mRNA sequence.
ACCESSION   BU430530
VERSION     BU430530.1  GI:22769017
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 358)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Eco RI site shown at the beginning of the sequence.
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
            cDNA Library Preparation: M.B. Soares Lab
            cDNA Library Arrayed by: M.B. Soares Lab
            DNA Sequencing by: M.B. Soares Lab
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: M13 Forward.
FEATURES             Location/Qualifiers
     source          1..358
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3066106"
                     /tissue_type="lymph"
                     /cell_type="germinal center B cells"
                     /cell_line="MGC85"
                     /lab_host="DH10B (LTI)"
                     /clone_lib="NIH_MGC_50"
                     /note="Vector: pT7T3-Pac; Site_1: NotI; Site_2: Eco RI;
```

Right column

```
          ,K.; Shingawa,A.; Yasunishi,A.; Yoshino,M.; Waterston,R.; Lander
          ,E.S.; Rogers,J.; Birney,E. and Hayashizaki,Y.
TITLE     Analysis of the mouse transcriptome based on functional annotation
          of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
MEDLINE   22354683
PUBMED    12466851
COMMENT   Contact: Yoshihide Hayashizaki
          Laboratory for Genome Exploration Research Group, RIKEN Genomic
          Sciences Center (GSC), Yokohama Institute
          The Institute of Physical and Chemical Research (RIKEN)
          1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
          Tel: 81-45-503-9222
          Fax: 81-45-503-9216
          Email: genome-res@gsc.riken.go.jp;
          URI:http://genome.gsc.riken.go.jp/
          Adachi,J.; Aizawa,K.; Akimura,T.; Arakawa,T.; Carninci,P.; Fukuda
          ,S.; Hashizume,W.; Hayashida,K.; Hirozane,T.; Hori,F.; Imotani,K.;
          Ishii,Y.; Itoh,M.; Kagawa,I.; Kawai,J.; Kojima,Y.; Kondo,S.; Konno
          ,H.; Koya,S.; Miyazaki,A.; Murata,M.; Nakamura,M.; Nomura,K.;
          Numazaki,R.; Ohno,M.; Ohsato,N.; Saito,R.; Sakazume,N.; Sano,H.;
          Sasaki,D.; Sato,K.; Shibata,K.; Shiraki,T.; Tagami,M.; Takeda,Y.;
          Waki,K.; Watahiki,A.; Muranatsu,M. and Hayashizaki,Y. Direct
          Submission
          Computational Analysis of Full-Length Mouse cDNAs Compared with
          Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
          Normalization and subraction of cap-trapper-selected cDNAs to
          prepare full-length cDNA libraries for rapid discovery of new
          genes. Genome Res. 10 (10), 1617-1630 (2000)
          RIKEN integrated sequence analysis (RISA) system--384-format
          sequencing Pipeline with 384 multicapillary sequencer. Genome Res.
          10 (11), 1757-1771 (2000)
          Computer-based methods for the mouse full-length cDNA
          encyclopedia: real-time sequence clustering for construction of a
          nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
          Encyclopedia Project of Genome Exploration Research Group in Riken
          Genomic Sciences Center and Genome Science Laboratory in RIKEN.
          Division of Experimental Animal Research in Riken contributed to
          prepare mouse tissues.
          Please visit our web site (http://genome.gsc.riken.go.jp) for
          further details.
FEATURES             Location/Qualifiers
     source          1..640
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="C430033A03"
                     /tissue_type="whole body"
                     /dev_stage="7 days embryo"
                     /clone_lib="RIKEN full-length enriched, 7 days embryo
                     whole body"
BASE COUNT   127 a   205 c   181 g   127 t
ORIGIN
Query Match
Best Local Similarity   20.7%;   Score 330.8;  DB 14;  Length 640;
Matches 422; Conservative   75.9%;  Pred. No. 5.7e-66;
         0; Mismatches 132; Indels  2; Gaps  1;

Qy     7 ATGGCGGGGGCCCGAGCGCTGGGGCCCCCCTTGCTCTCCTTGTGGCCTTCCTGCAGGGCG   66
         ||| ||||| || ||||| |||||||||||||||||||| |||||| |||||||| ||||
Db    87 ATGTGGGGGCGCCGACGGGTGGGGCCCCCTACTCTCGTTTCCGTGTTGCAGAGCGCGCGG  146

Qy    67 AGGCCCGCGTCTGGCCCCTCCCACCCCAGGAAACGTGACACTCTTCTCCCAGAACTTCAC  126
         |||||| || |   |||||||| |||| ||||| |||| | ||||||||||||||||| |
Db   147 AGGCCCGCGTCTAGCCCCACCCCAGGAAACGTGACACTCTTCTCCCAGAACTTCACTGTTTAC  206

Qy   127 CTGACAATGGCTCCCAGGGCTGGCAACCCCCAGGATGTGACCTATTTTGTGGCCATCAGA  186
         |||| |||||||||||||| | ||| |  |||||| ||||||||||||||||||  ||||
Db   207 CTGACATGGCTTCCGGGGGCTTGGAGCCCCCAAAATGACCTTATTTCTGTGACCTACCAA  266

Qy   187 GCTCTCCCACCCGTAGCAGGTGGGCCCGGAAGTGGAAGAGTGTGCGGGGAACCAAGGAGCTGC  246
```

Left column

```
          Constructed from size fractionated cytoplasmic mRNA
          (3.5-4.4kb). Directionally cloned. Cells provided by
          Louis M. Staudt, Ph.D. Library preparation by Maria de
          Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
BASE COUNT   67 a   112 c   108 g   70 t
ORIGIN
Query Match
Best Local Similarity   21.2%;   Score 338.2;  DB 13;  Length 358;
Matches 351; Conservative   98.6%;  Pred. No. 8.9e-68;
         0; Mismatches  4; Indels  1; Gaps  1;

Qy     5 CCATGGCGGGGGCGGGGCCGGCGGGGGCCGCTGGGGCCCCCCTTGCTCTCGCTGCTGCAGGGCGCGGCGGCCGCTCCAG   64
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3 CCAGGGCGGGGGCGGGGCCCGAGCGCTGGGGCCCCCCTTGCTCTCGCTGCTGCAGGGCGCGGCGGCCGCTCCAG   62

Qy    65 GGAGGCCCGCGTCTGGCCCCTCCCCAGAATGTGACGCTGCTCTCCCAGAACTTCAGCGTGT  124
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    63 GGAGGCCCGCGTCTGGCCCCTCCCCAGAATGTGACGCTGCTCTCCCAGAACTTCAGCGTGT  122

Qy   125 ACCTGACAATGGCTCCCAGGGCTGGCCAACCCCCAGGATGTGACCTATTTTGTGGCC-ATC  183
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   123 ACCTTACATGGCTCCCAGGGCTGGCCAACCCCCAGGATGTGACCTATTTTGTGGCCTATC  182

Qy   184 AGAGCTCTCCCACCCGTAGACGGTGGGCCGGGAGAGTGTGCGGGGAACCAAGGAGC  243
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   183 AGAGCTCTCCCACCCGTAGACGGTGGGCCGGGAGAGTGTGCGGGGAACCAAGGAGC  242

Qy   244 TGCTATGTTCTATGATGTGCCTGAAGAAACACGAGACCTGTACAACAAGTTCAAGGAGCGCG  303
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   244 TGCTATGTTCTATGATGTGCCTGAAGAAACACGAGACCTGTACAACAAGTTCAAGGAGCGCG  303

Qy   304 TGCTATGTTCTATGATGTGCCTGAAGAAACACGAGACCTGTACAACAAGTTCAAGGAGCGCG  303
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   303 TGCTATGTTCTATGATGTGCCTGAAGAAACACGAGACCTGTACAACAAGTTCAAGGAGCGCG  302

Qy   304 TGGGGACGGTTTCTCCCAAGTCCCCTCGGGGTGGGGGGGGACTCCGAATACCTGGAT  359
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   303 TGGGGACGGTTTCTCCCAAGTCCCCTTGGTGGTGGGAGCTCCGAATACCTGGAT  358
```

RESULT 11

```
LOCUS       AQ634533                  558 bp    DNA       linear    GSS 17-JUN-1999
DEFINITION  RPCI-11-478C4.TV RPCI-11 Homo sapiens genomic clone RPCI-11-478C4,
            genomic survey sequence.
ACCESSION   AQ634533
VERSION     AQ634533.1  GI:5097168
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 558)
AUTHORS     Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,
            J.C.
TITLE       Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
JOURNAL     Unpublished
COMMENT     Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@oregsen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
FEATURES             Location/Qualifiers
     source          1..558
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="CDB:7682319"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-478C4"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Seq_primer: T7 RPCI11 Human Male BAC Library"
                     Class: BAC ends.
BASE COUNT    113 a   156 c   180 g   109 t
ORIGIN

Query Match 18.3%; Score 292; DB 28; Length 558;
  Best Local Similarity 100.0%;
  Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy    267 --AGCTATATCAAAACCGGTTGGCGACCAGTGGGAGGCATTGTGCAGGTTATCAAGGCTCTGG 324
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    247 TATGTTCTATGATGTGCCTGGAGAAACAGACCTGTACAACAGAGTTCAAGGACGCGTGC   306

Qy    307 GGACGGGTTCTCCCAGTCCCAAGTCCCCTGGGGTGGGAGTCCGAATACTTGGATTACTTT 366
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    325 TGTGTCCCCCTGGATGTGCCTGAAGAAAACTGAACCGTACAGTTCAAAGGACGGAGTAC   384

Qy    367 TTGAAGTGGAGCGACGGCGGCCCCCCCACCTCGTCCTCGTGTCCCACCGGGGTGGAGTGG 426
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    385 AGGCACGTTCCGCACACGGGACGGTGCCCCCGTGCGAGGTCCCCGGTGACTCTGGAATACCC 444

Qy    427 CCAATGCCACAGTACCAGTGTCCCCTCGTCATGCCCCAGCGGATCTGAAGTATGAGGGTGG 486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    445 TTGAGGTGGAGCGTAGCCCCGCCCCCGCACCCTCGTTGTTCCTCATCCCCAGGAGGAATCCTTAAGG 504

Qy    487 CATTCTGGAAGGAGGGGGCCCCGGGGGAAAAACAAGACCCTATTCCAGTGCTCACTCCCCATGGCCAGC 546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    505 TCAACCCTACCTACCAGTGCCACCTTGCATGCCGTCTGGGAACTGAAGTACCAGGGTGG   564

Qy    547 CAGTCCAGATCACTCT 562
          ||||||||||||||||
Db    565 AATTCTGGAAGGAGGGGTCTGGGAAGACCCGTTCCTGACATCTCCTATGGCCAGC 624

Db    625 CAGTGCAGATTCCTCT 640
```

RESULT 12

```
LOCUS       BI104593                  620 bp    mRNA      linear    EST 26-JUN-2001
DEFINITION  602891434F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:5036481 5',
            mRNA sequence.
ACCESSION   BI104593
VERSION     BI104593.1  GI:14555486
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 620)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM1101 row: c  column: 10
            High quality sequence start: 5
            High quality sequence stop: 556.
FEATURES             Location/Qualifiers
     source          1..620
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C3ECH II"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5036481"
                     /clone_lib="NCI_CGAP_Lu29"
                     /tissue_type="spontaneous tumor, metastatic to mammary,
                     stem cell origin."
                     /lab_host="DH10B"
                     /note="Organ: lung, Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                     Library constructed by Life Technologies, Investigator
                     Gilbert Smith, NIH"
BASE COUNT    140 a   186 c   168 g   126 t
ORIGIN

Qy   1308 GGTTTCTCTGAAGAGCTCCCAGAAAGATAAACCTTCTCCTCCTCCTGGGGCCACTTC 1367
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGTTTCCTGGAAGAGCTCCCAGAAGATAAACCTTCTCCTCCTGGGGCCACCTTA     60

Qy   1368 CCACCCGGAGCGCGGAATCTGGTTCCCTGGGGGGGGGGGGACCCCCCAGTTTCTTCAGACACACTGACCTTC 1427
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CCACCCGGACGCGGAATCTGGTTCCCTGGGGGGGGGGGACCCCCCAGTTTTCTTCAGACACTGACCTTC   120

Qy   1428 TGCTGGGAAAGCAGCCCCTGAGGAGGAGGAGGAAGGAGGAAATCAGAAAATTCAGGACGAGCAGC 1487
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGCTGGGAAAGCAGCCCCTGAGGAGGAGGAGGAGGGAAATTGAGGACGAGC   180

Qy   1488 GATGCGGGCGCAGCTGTGGGCGGGGGGCTGAGGAGGCACCCAGAGGAGGCCCAGGAATTGAGGACGAGCAGC 1547
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GATGCGGGCGCAGCTGTGGGGGGGCTGGCTGAGAGCACCCAGAGGACCACCGGACATTG   240

Qy   1548 GGGCATTACATGGCCAGGTGAGGTGTGTCCCCCGACATCCCACCGGAATCTGATG 1599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGGCATTACATGGCCAGGTGAGGTGTGTCCCCCGACATCCCACCGGAATCTGATG 292
```

RESULT 13

```
Query Match       14.2%;     Score 227.2; DB 12; Length 620;
Best Local Similarity 70.3%;   Pred. No. 5.4e-42;
Matches 391; Conservative  0; Mismatches 153; Indels 12; Gaps  6;

Qy      7 ATGGCGGGGGGCGCCCGAGGCGGCTGGGGGCCCCGGCGGCCGGGCCGGGCCGGGGGCCGCCG-GCCGGCTCCAG 64
          |||||||||||||||||| || |||||||||||||||||||||||||||||||||||||||| |||||||||||
Db     16 ATGTGGCGGGGCGGACCGGTGTGCGCCCTACTCCTGTTTCCTGTTGTGCAGAGACCCCTAG 75

Qy     65 GGAGGCGCCCGTTCGGCCCCTTCCCAGAAGTGACGGCTGCTCTCCAGAACTTCAGCGTGT 124
          ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db     76 GAAGGCCCCGTCTAGCCCCTCACCCAGAAGCAGTGACACTTCCAGACTTCACTCGTTT 135

Qy    125 ACCTGACATGGCTCCCAGGGCTTGGCAACCCCCCAGGGATTGTGACCTATTTTTTGTGGCCATCA 184
          |||||||||| ||| ||||||||||||||||||||||||||||||||||||||||||||
Db    136 ACCTGACATGGCTTCCGGGGCTTGGGAGCCCCCCAAATGTGACCTATTTGTGTGACCTACC 195

Qy    185 GAGCTCTCCACCCGTTAGACGGTGGCGCGAAGTGGAAGAGAGTGTGCGGGGACCAAGGAGCT 244
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    196 AA--AGCTATATTCAAAAACCGGTTGGCGACCCAGTGGGACCATTGGTGCAGGTATCAAGGGCTCT 253

Qy    245 GCTATGTTCTATGAGTGTGCCTGAAGAAACAGGACCTGTACA-ACAAGTTCAAGGAGGCGCG 303
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    254 GGTGTGTCCCCTGATGTGCCTGAAGAAACTGGTACATCCAAGTTCAAAAGGACCTGAG 313

Qy    304 TGCGGAGGCGGTTTCTCCCAGTCCCCAAGTCCCCCTGGGGGGAGTCCCAAATACCTGGATTACC 363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    314 TACGAGGCAGCTTCCCCACAGAGGGTCCCACGGGGTGGGAGCCTCCCGGGTACCTGGAATACC 373

Qy    364 TTTTTGAAGTTGGGAGGCGGGGCCCCCCGACCACCTGTTCCTGTGGTGTCACCCACACGAGGAGGATCCTGA 423
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    374 TTTTTGACGTGGGAGGCTAGCCCCTGGGCTGCTGCTCCAGGTGGGAGTCCCGGGAATACC 433

Qy    424 GTGCCAAATGCCACGTACCAGCTGCCCCCCCACTGGATCTGAGTGTATGAGG 483
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    434 GGGTCA--TGCTACTACCAGAATGCCACTTGCAGGTGGCGTCACAAGTTGAACTACCAGG 491

Qy    484 TGGGCATTCTCTGGAAGGAGAGGGGGGGCCGGAAACAAGA-----CCCTATTTCAGTGCTCACTCCCCA 538
          |||||||||||||||||||||||||||||||||||||||||||||
Db    492 TGGGCAGTTCTTCGACAGGAGGGGCTCTGGGAAAGCAAGGACCCATGTCTCTGACACTCCCTA 551

Qy    539 TGGCCAGCCAGTCCAG 554
          ||||||||||||||||
Db    552 TGGCCACCAGTGCAG 567
```

BASE COUNT   109 a   155 c   135 g   97 t
ORIGIN

```
LOCUS       CB713061                   496 bp    mRNA    linear    EST 10-APR-2003
DEFINITION  AMGNNUC:NRHY5-00223-G11-A W Rat hypothalamus (10471) Rattus
            norvegicus cDNA clone nrhy5-00223-g11 5', mRNA sequence.
ACCESSION   CB713061
VERSION     CB713061.1  GI:29770209
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 496)
  AUTHORS   Amgen EST Program.
  TITLE     Amgen Rat EST Program
  JOURNAL   Unpublished
COMMENT     Contact: Dan Fitzpatrick
            Amgen, Inc
            One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
            Tel: 805 447-4881
            Plate: 00223  row: g  column: 11.
FEATURES             Location/Qualifiers
     source          1..496
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /clone="nrhy5-00223-g11"
                     /clone_lib="W Rat hypothalamus (10471)"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; W Rat
                     hypothalamus adult female Wistar rat avg. insert size 2.3
                     kb fraction 6 and 7"
```

RESULT 14

```
Query Match       13.5%;     Score 215.2; DB 14; Length 496;
Best Local Similarity 71.3%;   Pred. No. 3e-39;
Matches 325; Conservative  0; Mismatches 103; Indels 28; Gaps  2;

Qy      7 ATGGCGGGGGGCGCCCGAGGCGGCTGGGGGCCCCGGGGCCGGCCGGGGGCCGCCGGCCGGCTCCAGGG 66
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     55 ATGTGGCGGGCGGACCGGTGGCCGGTGGGCGCCCCTACTCCTGTTCTCCTGATGTGCAGAGA 114

Qy     67 AGGCCCGGTCTGGCCCCTTCCCAGGAATGTGAACGGTGCTCTCCAGAACTTCAGCGTGTAC 126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    115 AGGGCCCGGTCTAGCCCCTCACCCACAGAAAGGTGACACTTGCTCTCCAGGAGTTCACTGTTAC 174

Qy    127 CTGACATGGCTCCCAGGGCTTGGCAACCCCCCAGGGATGTGGACCTATTTTTGTGGCCATCAGA 186
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    175 CTGACATGGGCTCCAGGGCTAGGTAGCCCCCCGAAATGTGGACCTACTTGTGACCTACGAA 234

Qy    187 GCTCTCCACCCGTAGACGGTGGCGGGGAAGTGGAAGAGAGTGTGCGGGAACCAAGGAGCTGC 246
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    235 ----AGCTATAACAATCATTGGGCGAAGAGTAGAGCAATTGTGGGGGCATCAGAGTACTCTCG 289

Qy    247 TATGTTCTATGAGTGTGCCTGAAGAAAACAGGACCTGTACAACAACAAGTTCAAGGAGGCGCGTGC 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    290 TGTGTCCCATGAGTGTGCCTGAGAAACACAGGTATACATCCAAGTTCAAAAGGTCGGAGTAC 349

Qy    307 GGACGGTTTCTCCCAGTCCCCAAAGTCCCCTGGGGTGGACTCCCAAATACCTGGATTACTTTT 366
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    350 AGGCAGCTTCTGCCCATGTGAAGTCCCCATGGGGTGGAGTCC------------ 390

Qy    367 TTGAAGTGGAGCCGGGCCCCCCACCTGTCCTGTGTCTCACCGACGAGGAGGAGGATCTCGAGTG 426
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    391 ----AGTGGAGCTAGCCCCACCCACCCTGGTCTCTCCCAGAGAGGAGGAGATCCTAAATG 446

Qy    427 CCAATGCCACGTACCAGCTGCCCCCCCTGCATGCCC 462
          |||||||||||||||||||||||||||||||
Db    447 TCAACGCTACCTACCAGCTGCCACCCTGCATGCCTC 482
```

```
LOCUS       CB738101                   460 bp    mRNA    linear    EST 11-APR-2003
DEFINITION  AMGNNUC:MRPE3-00081-A5-A placenta embryo D17 (10379) Rattus
            norvegicus cDNA clone mrpe3-00081-a5 5', mRNA sequence.
ACCESSION   CB738101
VERSION     CB738101.1  GI:29805365
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 460)
  AUTHORS   Amgen EST Program.
  TITLE     Amgen Rat EST Program
  JOURNAL   Unpublished
COMMENT     Contact: Dan Fitzpatrick
            Amgen, Inc
            One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
            Tel: 805 447-4881
            Plate: 00081  row: a  column: 5.
FEATURES             Location/Qualifiers
     source          1..460
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /clone="mrpe3-00081-a5"
```

```
                    /tissue_type="placenta embryo"
                    /clone_lib="placenta embryo D17 (10379)"
                    /note="vector: pSPORT1; placenta embryo D17"
BASE COUNT     100 a     126 c     107 g      88 t     39 others
ORIGIN

Query Match          12.8%;   Score 204;  DB 14;  Length 460;
Best Local Similarity  74.0%;   Pred. No. 1.1e-36;
Matches  296;   Conservative   0;   Mismatches  99;   Indels   5;   Gaps   3;

Qy   433 CCACGTACCAGCTGCCCCCCCCTGCATGCCC---CACTGGATCTGAAGTATGAGGTGGCAT 489
             |||||  |||||||||| || |||  ||||      ||||| ||| ||| ||||| ||||
Db    39 CTACCTACCAGCTGCCCACCCCTGCATGCCTCGGTCACTGGAGACTGAAGTACCAGGTGGAGT 98

Qy   490 TCTGGAAGGAGGAGGGCCGGAAACAAGAGACCTATTTCCAGTCACTCCCCATGGCCAGCCAG 549
             ||||||||||||||||||||   |||||| |||||| ||||| |||||||||||||||||||
Db    99 TCTGGAAGGAGGAGGAACCTGAGAAGAACGACCCCGTTCCAGACACTCCATATGGCCAGCCAG 158

Qy   550 TCCAGATCACTCTTCAGCCAGCTGCCAGGCGAAACACCACTGCCTTCAGTGCCAGAACCATCT 609
             |||| |  |||| ||| ||||||  |||  |||||  ||| || ||||||||||||||||||
Db   159 TGAAGATTCCTTCCTCCAGCAAGATGCTCGTGGACGGCCACTGCCTTCAGCGCCGCCAGAACCGTCT 218

Qy   610 ACACGTTCAGTGTCCCGAAA-TACAGCAAGTTCTCTACACGCCACTGCCTTCAGCGCCCAGGAG 668
             ||| ||  ||  || |||     ||||   |||   |||| ||||  | || ||||| |||||
Db   219 ATACCTTAACCCACACNAAATTACAGCCTGTTCTCTGAGCCCAGCCTGATCTTCCTTANCG 278

Qy   669 GTCCCAGAAGCGGAACTCGGGCTTTCTCGTCGTGCCATCGCTTCTCGATACTGCTGTTAGTTA 728
             ||||| |||| |||| || |||||||| || || ||||| |||| || ||||| |||||||||
Db   279 GCCCCAAGGGCCACAAGGCTGTCCTCCTGCGCGGTGCCCCTTACTCTTGCTTCTACTGATAGCA 338

Qy   729 ATTGCCGCAGGGGGGTGTGATCTGGAAGGACACCCTCATGGGGGAACCCTGGTTTCAGCGGGCA 788
             |||| |  | ||| ||| |||   |||||| |||| ||| ||||| ||| || ||||||||||
Db   339 ACGGTGCACGCGGTACGGCATGGAGAAAATGATGGGAA-CCCTGGTTTCAGTGGGGTG 397

Qy   789 AAGAATGCCACGGGCCCCTGGACTTTTCTCGGGACACACACACC 828
             ||| |  ||||| ||| |||| |||||| || ||||| |||||||
Db   398 AAGATCCCCCANGCACTGGACTTTTCTGANNAACAGAGACAC 437
```

```
FEATURES             Location/Qualifiers
     source          1..506
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                    /clone="Plate=989 Col=2 Row=P"
                    /sex="male"
                    /clone_lib="RPCI-11 Human Male BAC Library"
                    /note="vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;
                    Male blood DNA was isolated from one randomly chosen donor
                    and partially digested with a combination of EcoRI and
                    EcoRI Methylase. Size selected DNA was cloned into the
                    pBACe3.6 vector at EcoRI sites"
BASE COUNT     111 a     153 c     144 t      97 g      1 others
ORIGIN

Query Match          12.6%;   Score 201.2;  DB 28;  Length 506;
Best Local Similarity  80.9%;   Pred. No. 5.2e-36;
Matches  258;   Conservative   0;   Mismatches  59;   Indels   2;   Gaps   2;

Qy   969 CTTGCAGAGGACGCAGAGAGGAGGAGGAGAGGAGGATGAGGAGGAGGACAGAAGATGGGCTTCAGCTTCAG 1028
             ||||| || ||| ||| || ||||||  || || |||||||| ||| ||| ||| |||| ||||||||||
Db   322 CTTTGAATGCACAAAAGAGTAGGAGTAATG-GGAAGACACATCGAGATGGTGTTAAGTACCAA 264

Qy  1029 CCCTACATTGAACCACCTTCTTTCCTGGGGCAAGAGACCACCAGGCTCCAGGGCCACTCCGGAG 1088
             ||||||  | |||||||  ||||||| |||| || ||  ||| ||||| || ||||| |||||
Db   263 CCGTACATTGAACCTCGTTTTATCATGGGG-AAGAGCTCAGGGCACTACGAGAG 205

Qy  1089 GCTTGGGTGGGGTTGGGACTTCAGGGAGGCCCCAGGGGTCCTCGTCTCCCAAAGGCAAGGCTCCTCT 1148
             ||||| || |||| |||| || ||| ||||| || ||||| ||||||| ||| ||||| ||| || |
Db   204 TCTGGTGGGGTTGGCTCAGGGAGGCACGGTGCTCATTGGTCCCCAAGAGAAGAGGTTCATCT 145

Qy  1149 GCTTGGGATTCCTTCAGACAGCAGGAAGTTGGCCCAGCAGAGTGCACTCCTCCTGGGGACAGGGCT 1208
             |||| |||||| |||||| ||||| || ||||| ||||| || ||| | ||| ||||| ||||||
Db   144 GCGTGGGATTCACAGCACCAGAGAATTTGACCTCCTTCCTGGGGACAGGGCC 85

Qy  1209 GGGTCCTTCCTGGCTATTTGGTGAGAAGGGGCCCAGGGCCCGGGGCTGGGGGATGGGGCAC 1268
             ||||| | | || |||||| |||| | |||| || |||| ||| ||| ||||| ||||||
Db    84 GGGTCCCTCTGGCTATTTGGTAAGAAGGGGACAGTCAGAGGGGGCCGGGGTGGGANTGGGCAC 25

Qy  1269 CAAGAATCTCTCCCACCAC 1287
             |||||||| | ||||||||
Db    24 CAAGAATCTAACACTCCAC 6
```

```
RESULT 15
LOCUS       AQ772013/c
DEFINITION  AQ772013  HS_5413_B2_H01_SP6E RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=989 Col=2 Row=P, genomic survey sequence.
ACCESSION   AQ772013
VERSION     AQ772013.1  GI:5651741
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 506)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  PUBMED    9380589
  MEDLINE   10449764
  COMMENT   Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering.bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
```

Search completed: September 17, 2003, 22:13:24
Job time : 2379.81 secs